IMERG Quality Index
George Huffman
9 November 2017, Rev. 1 February 2018

There have been several requests from users recently for a "simple" quality index (QI) to give some guidance on when they should most trust the Integrated Multi-satellitE Retrievals for GPM (IMERG). While the goal is reasonable, there is no agreement about how this quantity should be defined. After some discussion within the team, two distinctly different quality indices were chosen for the half-hourly and monthly data fields (QIh and QIm, respectively) for implementation in Version 05. It is a matter of investigation to determine if users find these insightful, or if different quality indices should be developed for future releases.

QIh: Quality Index for Half-Hourly Data

At the half-hourly scale, the best metric is some measure of the relative skill that might be expected from the fluctuating mix of different passive microwave- and infrared-based precipitation estimates. The Kalman smoother used in IMERG (and originated in the CPC KF-CMORPH algorithm, Joyce et al. 2011) routinely updates estimates of correlation between GMI and each of the other satellite estimates in separate coarse land and ocean blocks across the entire latitude band 60°N-S, and then uses these correlation coefficients (squared) to provide weights for use in the combination of forward-propagated passive microwave, backward-propagated passive microwave, and current-time (nominally taken as the +30 minute field) infrared precipitation estimates. Specifically, the correlations are computed for each half-hour forward and backward "time step" away from the current half hour, separately for imager and sounder estimates. Because there is no formalism for computing an overall correlation for the combined estimate, one approach is provided here.

The usual approach is to compute the RMS of a combined estimate (σ_t) in terms of the individual RMS estimates (σ_a and σ_b), which is given as

$$\sigma_t = \frac{\sigma_a \sigma_b}{\sqrt{\sigma_a^2+\sigma_b^2}} \qquad 1$$

The KF-CMORPH Kalman smoother uses squared correlation coefficient (c^2) in place of $1/\sigma^2$ in the weighting of the input precipitation estimates, so substituting $1/c$ for σ in (1) and simplifying,

$$c_t = \sqrt{c_a^2 + c_b^2}\,, \qquad 2$$

where c_a and c_b are individual correlation coefficients for estimates a and b, and c_t is the estimated correlation coefficient for the combination of estimates a and b.

This formulation has the advantage of producing correlation coefficients higher than the individual input terms, highest when c_a and c_b are equal, and declining to c_a as c_b goes to zero (and vice-versa). However, for both c's close to 1, the resulting c_t can exceed 1 and be as high as

1.414 (square root of 2). One solution to this quandary is to introduce a variance-stabilizing transformation. One simple choice is the Fisher (1915) z statistic

$$z = \text{arctanh}(c)\,, \qquad 3$$

where c is a correlation value. The transformed value z takes on large values as c approaches 1 (or -1), so transforming to z, performing calculations with z, and back-transforming avoids problems around 1. Substituting z for c in (2),

$$c_t = tanh\left(\sqrt{arctanh^2(c_a) + arctanh^2(c_b)}\right) \qquad 4$$

the ordering remains and it gracefully approaches 1. Formally, the Fisher transformation requires that the two variables being correlated follow a bivariate normal distribution. While this is not true for precipitation, we adopt this approach as a first approximation to computing the correlation coefficient of the combined precipitation estimate because its use as a quality index seems reasonable and useful. In the case of three input correlation coefficients, the equation simply extends to three terms on the right-hand side. The units are non-dimensional correlation coefficients. The equation as applied to IMERG is

$$c_t = tanh\left(\sqrt{arctanh^2(c_{fp}) + arctanh^2(c_{bp}) + arctanh^2(c_{ir})}\right)$$

where c_{fp} is the forward propagated microwave estimate, c_{bp} is the backward propagated microwave estimate, and c_{ir} is the IR estimate. Note that IR estimates are only included when the microwave propagation is beyond +/-90 minutes from the current half-hour.

There is one additional issue: we lack the zero half-hour correlation of each constellation member to the GMI for computational reasons in the current implementation of IMERG and need an approximate value. Lacking strong justification for alternatives, we chose to set c_t = 1 when the current half-hour microwave estimate is present.

The c_t thus defined is adopted as QIh. The next version of this approach should revisit this choice.

QIm: Quality Index for Monthly Data

At the monthly scale, a relatively well-founded metric exists for random error, based on Huffman's (1997) analysis of sampling error for a particular data source for a month. The general form of the relationship is

$$\sigma_r^2 = \frac{\bar{r}^2}{N_I}\left(\frac{H}{p} - 1\right), \qquad 5$$

where σ_r is random error, $\bar{r}$ is the time-average of the precipitation rate (originally labeled "rain rate") samples, N_I is the number of independent samples in $\bar{r}$, H is the non-dimensional second moment of the probability distribution of the precipitation rates, and p is the frequency of all nonzero precipitation. Huffman (1997) proceeds to simplify (5) to the approximate expression

$$\sigma_r^2 \cong \frac{H}{I}\frac{(\bar{r}+S)}{N}\left[24+49\sqrt{\bar{r}}\right], \qquad 6$$

where $\bar{r}$ and N are available for each grid box in the monthly estimate, I is a multiplicative constant expressing the fraction of N that is "independent", and H/I and S are global constants that are approximated with validation data for each sensor type. This relationship is simple enough that it can be inverted for N. When all the constants are set for the gauge analysis, but the $\bar{r}$ and σ_r^2 used are the final satellite-gauge precipitation estimate and random error variance,

$$N \cong \left(\frac{H}{I}\right)_g \frac{(\bar{r}+S_g)}{\sigma_r^2}\left[24+49\sqrt{\bar{r}}\right], \qquad 7$$

and this special N is defined as the equivalent number of gauges. Following Huffman (1997), the interpretation is that this is the approximate number of gauges required to produce the estimated random error, given the estimated precipitation. The units are gauges per area, and in the current implementation the area is carried as 2.5°x2.5° of latitude/longitude, even though IMERG is computed on a much finer scale, in order to facilitate interpretation in large-error regions.

N, the equivalent number of gauges, is adopted as QIm. Note that N is dominated by the number of gauges except where gauges are sparse.

Examples

An example of QIh for the IMERG Final Run is shown in Fig. 1. The thin strips of lower QIh are microwave estimates that have longer propagation times between current half-hour microwave swaths. Blockiness is due to the regional variations caused by the coarse resolution and land-ocean separation in the background correlation statistics. Low values at high latitudes are due to two factors. First, microwave estimates are masked out over snowy/icy surfaces, so these regions only have microwave-adjusted IR-based estimates, which have inherently lower correlations. Second, the microwave adjustment to the IR depends on adjustments interpolated from surrounding areas to the areas where microwave estimates have been screened out due to snowy/icy surface. As noted before, grid boxes carrying current-half-hour data from passive microwave input are presently given values of 1, even though the actual correlation should be somewhat lower.

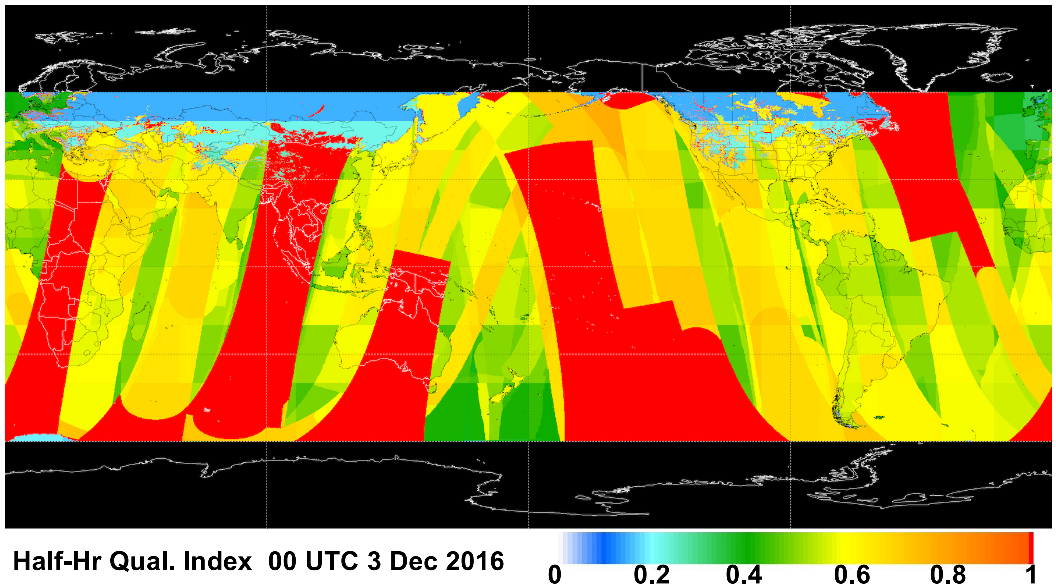

Fig. 1. QIh (computed as a composite correlation) for the half-hourly IMERG Final Run for the period 0000-0030 UTC on 3 December 2016. Blacked-out areas lack data. [Courtesy D. Bolvin (SSAI; GSFC)]

An example of QIm for the IMERG Final Run is shown in Fig. 2. [Recall that only the Final Run has monthly data as a native product.] Over oceans, the equivalent gauges metric largely tames the variation of random error with precipitation rate as the sampling by the satellite estimates is relatively uniform. Over land, QIm largely reflects the distribution of precipitation gauges, except it has the lower limit of the satellite equivalent gauges (similar to the values over ocean) where gauges are extremely sparse. The QIm values outside the morphing region (60·N-S) reflect relatively sparse gauges (over snowy/icy land) and passive microwave sampling over ice-free ocean.

<u>Advice on Using the Quality Index</u>

QIh is still a work in progress, so advice on its use is necessarily preliminary. Early testing by the developers using the Multi-Radar Multi-Sensor (MRMS) data over CONUS seems to show that the (highly sparse) 0-0.2 range of QIh values is clearly different, while the top value of QIh = 1 is better in correlation and scatter diagrams, and most of the metrics are similar to the middle range of QIh values. One interesting difference is that the 0.2-0.4 range in QIh has much higher fractional coverage (sum of hits and misses) by precipitation and 0.8-1 is somewhat higher. The scatter diagrams for a large range of middle QIh values show more counts in both the low and middle-high IMERG values, although they lack the highest values that IMERG has for QIh = 1. The scatter diagram for QIh in the range 0.2-0.4 has the most low precipitation values and the fewest mid-to-high.

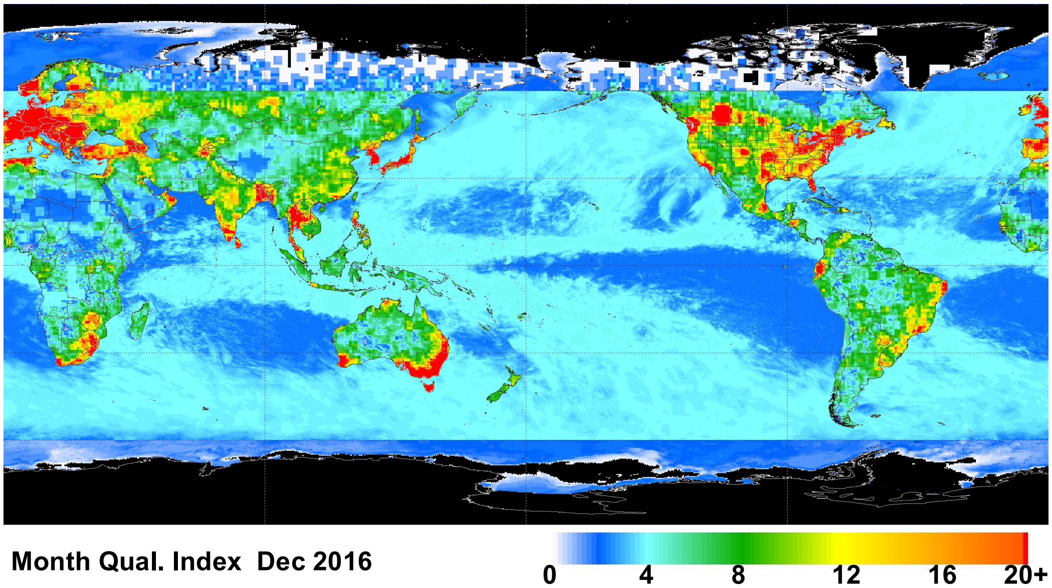

Fig. 2. QIm (computed as equivalent gauges per 2.5°x2.5° lat./lon. box) for the Monthly IMERG Final Run for December 2016. Blacked-out areas lack data. [Courtesy D. Bolvin (SSAI; GSFC)]

As a three-class "stoplight" statement, the preliminary advice is therefore:
0-0.3 = "red" significant IR contribution with high uncertainty
0.3-0.9 = "yellow" the mid-range (~70% of the values) is morphed microwave
0.9-1 = "green" this is the current half-hour microwave swath data

Note that the numbers of cases around 0.3 and 0.9 are extremely low, so the exact choice of threshold is not critical.

QIm has a longer history at the 2.5° scale, but is relatively new for the 0.1° scale. Based on an experience with regions with different QIm values, the preliminary advice on "stoplight" values is:
0-2 = "red" equivalent to the gauge coverage in regions such as central Africa, where the lack of data in a gauge-only analysis a critical problem
2-10 = "yellow" the mid-range has enough gauge data to ensure reasonable bias adjustment, but still require interpolation to fill in gaps several grid boxes wide between stations more or less routinely
10+ = "green" these are developed areas with good-to-excellent gauge networks

<u>References</u>

Fisher, R.A., 1915: Frequency Distribution of the Values of the Correlation Coefficient in Samples of an Indefinitely Large Population. *Biometrika*, Biometrika Trust, **10**(4), 507–521. *doi:10.2307/2331838*

Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global,

Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. *J. Hydrometeor.*, **8**, 38-55. *doi:10.1175/JHM560.1*
Joyce, R.J., P. Xie, J.E. Janowiak, 2011: Kalman Filter Based CMORPH. *J. Hydrometeor.*, **12**, 1547-1563. *doi:10.1175/JHM-D-11-022.1*